Transactions of the Institute of Measurement and Control 30, 2 (2008) pp. 173–199

Difference equation sensor characterization algorithms for two-thermocouple probes

Seán McLoone[1], Peter Hung[2], George Irwin[2] and Robert Kee[2]

[1]Department of Electronic Engineering, National University of Ireland Maynooth (NUIM), Maynooth, Co. Kildare, Ireland
[2]Virtual Engineering Centre, Queen's University Belfast, Belfast BT9 5HN, Northern Ireland

The characterization of thermocouple sensors for temperature measurement in variable flow environments is a challenging problem. In this paper, novel difference equation-based algorithms are presented that allow *in situ* characterization of temperature measurement probes consisting of two-thermocouple sensors with differing time constants. Linear and non-linear least squares formulations of the characterization problem are introduced and compared in terms of their computational complexity, robustness to noise and statistical properties. With the aid of this analysis, least squares optimization procedures that yield unbiased estimates are identified. The main contribution of the paper is the development of a linear two-parameter generalized total least squares formulation of the sensor characterization problem. Monte-Carlo simulation results are used to support the analysis.

Key words: sensor characterization; soft sensing; two-thermocouple probe.

Nomenclature

$T_g(t)$	True gas temperature (°C)
$T_m(t)$	Measured gas temperature (°C)
$T_1(t)$	Temperature measured by thermocouple 1 (°C)

Address for correspondence: Seán McLoone, Department of Electronic Engineering, National University of Ireland Maynooth (NUIM), Maynooth, Co. Kildare, Ireland.
E-mail: s.mcloone@ieee.org

© 2008 The Institute of Measurement and Control 10.1177/0142331208092148

$T_2(t)$	Temperature measured by thermocouple 2 (°C)
T_1^k	Temperature measured by thermocouple 1 at kth sample instant (°C)
T_2^k	Temperature measured by thermocouple 2 at kth sample instant (°C)
T_g^k	True gas temperature at kth sample instant (°C)
T_s	Sampling period (s)
τ_1	Time constant of thermocouple 1 (s)
τ_2	Time constant of thermocouple 2 (s)
α	Ratio of time constants, $\alpha = \tau_1/\tau_2$, $\alpha < 1$ by definition
ν_1	Variance of noise on measurements from thermocouple 1
ν_2	Variance of noise on measurements from thermocouple 2
ϕ	Ratio of the noise variances, $\phi = \nu_1/\nu_2$
a_1, b_1	Parameters of the discrete-time model for thermocouple 1
a_2, b_2	Parameters of the discrete-time model for thermocouple 2
β	Ratio of b_2 and b_1, $\beta = b_2/b_1$
ΔT_i^k	Change in measured temperature from the $(k-1)$th to the kth sample instant for the ith thermocouple, ie, $\Delta T_i^k = T_i^k - T_j^k$ (°C)
ΔT_{ij}^k	Difference in temperature measured by the ith and jth (°C) thermocouple at the kth sample instant, ie, $\Delta T_{ij}^k = T_i^k - T_j^k$
$x_k, y_k, \boldsymbol{\theta}$	Generic regression, output vector and parameter vector in $y_k = \mathbf{x}_k^{\mathrm{T}}\boldsymbol{\theta}$
X, y	Matrix form of equation $\mathrm{y} = \mathbf{X}\boldsymbol{\theta}$
TTP	Two-thermocouple probe
LS	Least squares
TLS	Total least squares
GTLS	Generalized total least squares
SVD	Singular value decomposition

1. Introduction

Temperature measurement can be performed by exploiting a variety of temperature-related material properties such as variation in resistivity (resistance temperature detectors and thermistors), variation in volume (mercury thermometer), differential expansion (bimetal strip), dissimilar metal junction potential (thermocouple) and radiated heat/light (optical pyrometry). Of these approaches the thermocouple is the most widely used in commercial applications because of its robustness and low cost of manufacture. The main drawback is that thermal inertia thermocouple response times are relatively slow. Consequently, they are only appropriate in applications where the temperature changes relatively slowly (<1 Hz).

The frequency range of thermocouples can be extended by as much as 10 to 100 times by using software-based compensation schemes. Most compensation schemes rely on having a dynamic model of the sensor whose parameters are determined through an initial calibration procedure (sensor characterization). In the case of

thermocouples in a gas or liquid flow, a first-order model with time constant, τ, is generally assumed (Kee *et al*., 1999).

$$\tau \frac{dT_m}{dt} + T_m = T_g \tag{1}$$

A common *a priori* calibration technique is to heat the thermocouple by passing a current through the thermocouple wire and then allow it to cool down in the environment in which it is being used. The time constant, τ, can then be estimated from the resulting cooling curve. The difficulty with this approach is that τ is strongly dependent on the physical and mechanical properties of the thermocouple and its environment, and therefore *a priori* characterization is only applicable when these conditions are invariant during sensor operation. In many situations, such as the measurement of temperature in a varying flow environment, this is not the case. Here the time constant of a thermocouple is related to its diameter according to the equation

$$\tau = kd^{2-m}v^{-m} \tag{2}$$

where k and m are constants (approximately) arising from thermodynamic considerations, d is the diameter of the thermocouple wire and v is the velocity of the gas/liquid medium in which it is placed.

Using a probe consisting of two thermocouples (Figure 1) with different time constants, it is possible to identify both time constants *in situ* and subsequently reconstruct the input temperature. The underlying assumption is that due to their close proximity, both thermocouples are subject to the same environmental conditions, hence the same temperature $T_g(t)$ and medium velocity v. Under these circumstances, it follows from Equation (2) that the ratio of the time constants, given by

$$\alpha = \frac{\tau_1}{\tau_2} = \frac{kd_1^{2-m}v^{-m}}{kd_2^{2-m}v^{-m}} = \left(\frac{d_1}{d_2}\right)^{2-m}, \quad \alpha < 1 \tag{3}$$

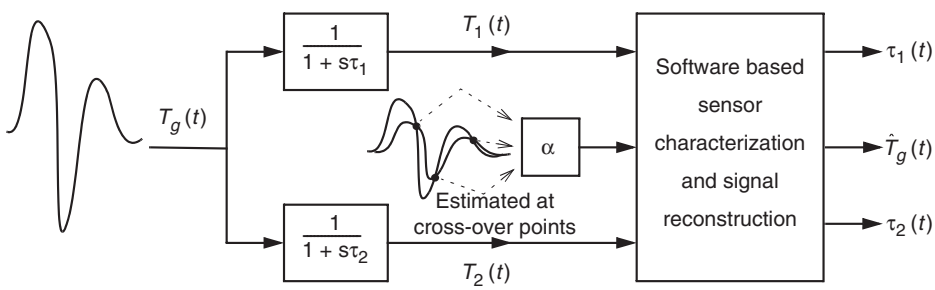

Figure 1 TTP-based *in situ* thermocouple characterization

is a function of thermocouple geometry only, and therefore approximately invariant. Here subscripts 1 and 2 are used to distinguish between the two thermocouples and the corresponding thermocouple models are given by

$$\tau_1 \frac{dT_1}{dt} + T_1(t) = T_g(t) \tag{4}$$

$$\tau_2 \frac{dT_2}{dt} + T_2(t) = T_g(t) \tag{5}$$

respectively. Assuming knowledge of α, *in situ* instantaneous estimates of the time constants are given by

$$\tau_2(t) = \frac{T_1(t) - T_2(t)}{\dot{T}_2(t) - \alpha \dot{T}_1(t)}, \quad \tau_1(t) = \alpha \tau_2(t) \tag{6}$$

Indeed $T_g(t)$ can be estimated directly without having to determine τ_1 and τ_2 explicitly using

$$T_g(t) = \frac{T_1(t)\dot{T}_2(t) - \alpha \dot{T}_1(t) T_2(t)}{\dot{T}_2(t) - \alpha \dot{T}_1(t)} \tag{7}$$

It is believed that this approach was first suggested by Pfreim (1936) and subsequently rediscovered by Strahle and Muthukrishman (1976) and Cambray (1986). Strahle and Muthukrishman (1976) developed a procedure for estimating the time constants *in situ* by analysing the cross and auto power spectra of the probe signals, while Cambray (1986) exploited the invariance of the time constant ratio, α, to reduce the problem to one of *a priori* ratio estimation. In particular, Cambray noted that since

$$\alpha \frac{\dot{T}_1(t)}{\dot{T}_2(t)} = \frac{T_g(t) - T_1(t)}{T_g(t) - T_2(t)} \tag{8}$$

α can be computed as the ratio of the slopes of the temperature responses $T_1(t)$ and $T_2(t)$ at crossover points, ie, when $T_1(t) = T_2(t)$

$$\alpha = \frac{\dot{T}_2(t)}{\dot{T}_1(t)} \tag{9}$$

In recent years, researchers at the NASA–Lewis Research Centre for jet engine studies in the USA, the Combustion Laboratory at the Nagoya Institute of Technology in Japan and the Internal Combustion Engine Research Group at Queen's University Belfast in Northern Ireland have developed these concepts further to produce more robust two-thermocouple probe (TTP) *in situ* characterization and signal reconstruction algorithms. Tagawa and Ohta (1997), Tagawa *et al*. (1998), Kee *et al*. (1999), and O'Reilly *et al*. (2001) developed time-domain methods, while Forney and Fralick

(1994, 1995a, 1995b) and Tagawa *et al*. (2003) transformed the problem to the frequency domain (using the FFT), thereby avoiding signal derivative estimation.

While these TTP methods are a major step forward they have a number of weaknesses:

- Algorithm performance deteriorates rapidly (and in some instances catastrophically) as the signal-to-noise ratio increases. Although the frequency domain methods are better able to deal with noise, their performance also deteriorates rapidly. Furthermore, they are severely restricted with respect to the class of signals to which they apply.
- None of the existing methods has been developed within a statistical framework and, as such, are not guaranteed to produce unbiased estimates in the presence of measurement noise.
- With the exception of Tagawa and Ohta (1997) and Tagawa *et al*. (1998, 2003), the TTP schemes rely on the invariant time constant ratio (α) property identified by Cambray (1986) and thus depend on accurate *a priori* estimation of this parameter. This drawback limits the applicability of these methods.
- Many of the methods suffer from numerical issues such as singularities and sensitivity to measurement offsets.

Recently the authors proposed a novel discrete-time formulation of the TTP sensor characterization problem that does not require the invariant α assumption and allows the problem to be cast as a linear input–output system identification problem whose parameters are algebraically related to the desired time constants [subject to a zero-order-hold (ZOH) approximation, Hung *et al*., 2002, 2003]. The significance of this new approach is that it allows the introduction of a statistical framework and the application of a vast body of theory and techniques from the field of linear system identification (Ljung, 1999). For example, this initial work has highlighted that the estimation task arising in TTP characterization falls into the Error-In-Variable (EIV) class of system identification problems and consequently conventional Least Squares (LS) parameter estimation methods will produce biased estimates. In Hung *et al*. (2003), Total Least Squares (TLS) is suggested as a way of obtaining unbiased estimates when the noise on the thermocouple measurements is white and of equal variance.

In this paper, the discrete-time approach to TTP characterization is developed in full. Various linear and non-linear formulations are introduced and compared in terms of their computational complexity, robustness and statistical properties. With the aid of an analysis of the statistical properties of these formulations, LS optimization procedures, which yield unbiased parameter estimates, are identified.

Two simulated temperature variation test cases will be used throughout the paper to demonstrate the properties of the various algorithms developed. The first is a a two-tone sinusoidal signal defined as

$$T_g(t) = 50 + 15\sin(20\pi t) + 5\sin\left(34\pi t + \frac{\pi}{3}\right) \tag{10}$$

while the second is a broadband signal consisting of band-limited white noise (0–1000 Hz) and sinusoidal tones at 15, 20, 25, 30 and 35 Hz. Both signals have a strong sinusoidal component reflecting the cyclic nature of temperature variation observed in many practical applications (eg, engine exhaust gas temperature measurement, Kee *et al*., 1999). The latter includes band-limited noise to represent random fluctuations under turbulent flow conditions. Samples of these signals and their frequency spectra are depicted in Figures 2 and 3 along with the responses obtained from 0.02-s and 0.1-s time constant thermocouples.

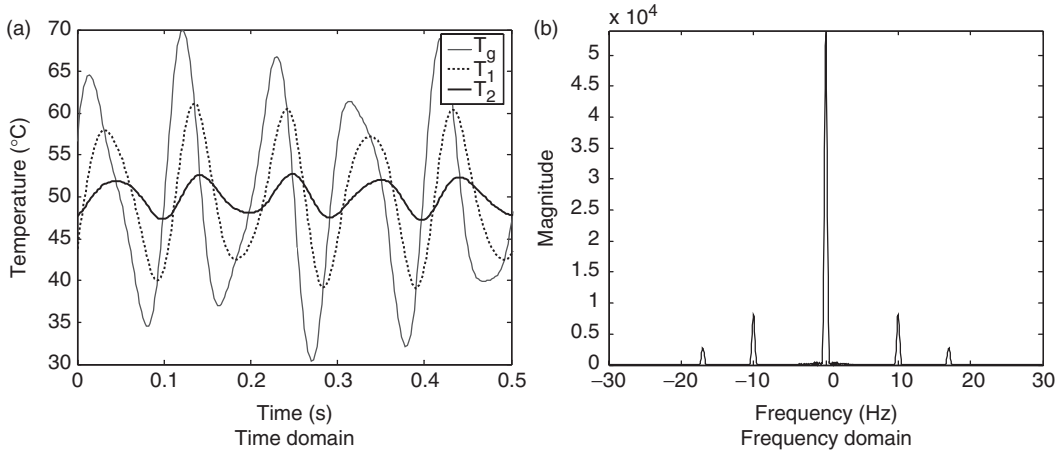

Figure 2 Two-tone simulated gas test signal: (a) sample of signals T_g, T_1 and T_2; (b) frequency spectrum of T_g

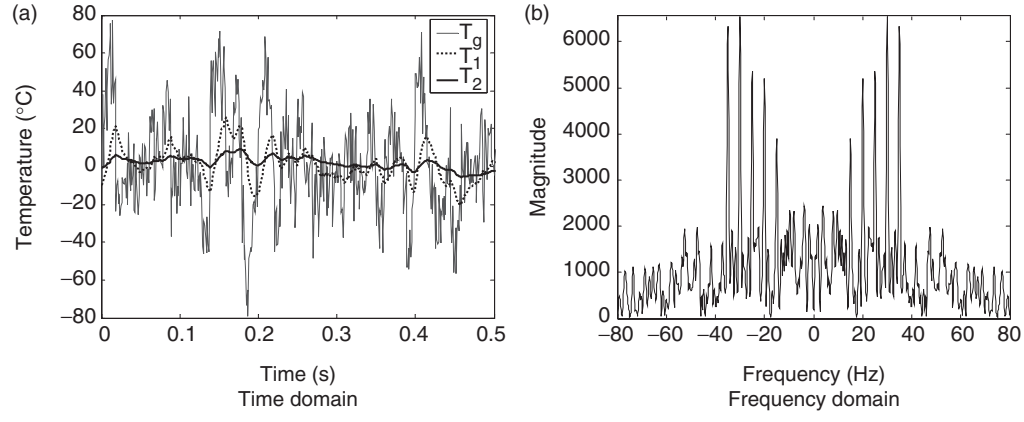

Figure 3 Multi-tone random simulated gas test signal: (a) sample of signals T_g, T_1 and T_2; (b) frequency spectrum of T_g

The simulated signals are sampled at 500 Hz and 1 kHz, respectively, yielding 1000 sample data sets. Zero mean normally distributed random numbers are added to the samples to simulate white measurement noise. The noise level defined as

$$\text{Noise level} = \frac{std(\text{noise})}{std(\text{signal})} \times 100 \tag{11}$$

is used to quantify the amount of noise introduced. Here, *std* is the standard deviation operator.

The remainder of the paper is structured as follows. Section 2 introduces some preliminaries on LS optimization focusing on the problem of biased estimation when noise is present on regressor variables. The TTP discrete-time characterization approach is then outlined in Section 3 and the novel characterization formulations presented in Section 4. Monte-Carlo simulation results comparing the different methods are given in Section 5 and finally conclusions are presented in Section 6.

2. Least squares (LS) preliminaries

2.1 Singular value decomposition (SVD)

The SVD of a $n \times m$ matrix $\mathbf{A}$ is defined as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}} = \sum_{i=1}^{p} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}} \tag{12}$$

where

$$\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_n], \quad \mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_m], \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{13}$$

$$\mathbf{S} = diag(\sigma_1, \ldots, \sigma_p), \quad p = \min\{m, n\} \text{ and } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0 \tag{14}$$

$\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices (ie, $\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_n$, $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_m$) referred to as the left and right singular matrices, respectively, and $\mathbf{\Sigma}$ is the matrix of singular values. Scalar σ_i is the ith singular value of $\mathbf{A}$ and vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are the corresponding left and right singular vectors. If matrix $\mathbf{A}$ is of rank r and $r < p$, then it follows that only the first r singular values of $\mathbf{A}$ are non-zero and the SVD expansion can be written as

$$\mathbf{A} = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^{\mathrm{T}} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}} \tag{15}$$

where

$$\mathbf{U}_r = [\mathbf{u}_1 \ldots \mathbf{u}_r], \quad \mathbf{V}_r = [\mathbf{v}_1 \ldots \mathbf{v}_r], \quad \mathbf{\Sigma}_r = diag(\sigma_1, \ldots, \sigma_r) \tag{16}$$

Equation (15), also known as the dyadic decomposition of matrix $\mathbf{A}$, is the expansion of $\mathbf{A}$ as an ordered sum of rank one matrices, $\sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}}$. Significantly, this decomposition is such that the closest rank q matrix, $\mathbf{A}_q^*$ to $\mathbf{A}$ is given by

$$\mathbf{A}_q^* = \sum_{i=1}^{q} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}} \tag{17}$$

where closeness is measured in terms of the Frobenius norm of the difference between $\mathbf{A}_q$ to $\mathbf{A}$, ie, $\|\mathbf{A} - \mathbf{A}_q\|_F$. By noting that $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{V}\,\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}}$ and $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{U}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{U}^{\mathrm{T}}$ it can be seen that the singular values are the square roots of the eigenvalues of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ and $\mathbf{A}\mathbf{A}^{\mathrm{T}}$, and $\mathbf{v}_i$ and $\mathbf{u}_i$ are the corresponding eigenvectors.

2.2 Least Squares (LS)

In the conventional LS formulation, we have a linear model of the form

$$y_k = \mathbf{x}_k^{\mathrm{T}}\boldsymbol{\theta} \tag{18}$$

where $\mathbf{x}_k$ is the $p \times 1$ regression vector, y_k is a scalar output and $\boldsymbol{\theta}$ is the $p \times 1$ vector of unknown parameters. For a set of n samples, the regression matrix and corresponding output vector can be defined as

$$\mathbf{X} = [\mathbf{x}_1 \ldots \mathbf{x}_n]^{\mathrm{T}} \text{ and } \mathbf{y} = [y_1 \ldots y_n]^{\mathrm{T}} \tag{19}$$

leading to the matrix equation

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} \tag{20}$$

Here $\mathbf{X}$ is an $n \times p$ matrix and $\mathbf{y}$ is an $n \times 1$ vector. The LS estimate of $\boldsymbol{\theta}$ is then given by

$$\boldsymbol{\theta}_{\mathrm{LS}} = [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} = \mathbf{X}^{\dagger}\mathbf{y} \tag{21}$$

Because of numerical issues, the pseudo-inverse ($\mathbf{X}^{\dagger}$) is seldom computed directly. Instead robust procedures such as SVD are employed. Thus, given the SVD of $\mathbf{X}$,

$$\mathbf{X} = \mathbf{U}_r \Sigma_r \mathbf{V}_r^{\mathrm{T}}, \quad r = rank(\mathbf{X}) \tag{22}$$

as defined in (15), the pseudo-inverse can be computed as

$$\mathbf{X}^{\dagger} = \mathbf{V}_r \Sigma_r^{-1} \mathbf{U}_r^{\mathrm{T}} \tag{23}$$

2.3 Statistical properties of the LS estimate

To evaluate the statistical properties of the LS solution in the presence of noise, it is useful to express the solution in terms of the sample covariance matrix $\mathbf{R}_n$,

$$\mathbf{R}_n = \frac{\mathbf{X}^{\mathrm{T}}\mathbf{X}}{n} \tag{24}$$

and the sample cross-correlation vector, $\mathbf{p}_n$,

$$\mathbf{p}_n = \frac{\mathbf{X}^{\mathrm{T}}\mathbf{y}}{n} \tag{25}$$

By definition, as the number of data points, n, tends to infinity we obtain the true correlation matrix and cross-correlation vector, ie,

$$\mathbf{R} = \lim_{n\to\infty} \mathbf{R}_n = E[\mathbf{x}_k \mathbf{x}_k^{\mathrm{T}}] \tag{26}$$

$$\mathbf{p} = \lim_{n\to\infty} \mathbf{p}_n = E[y_k \mathbf{x}_k^{\mathrm{T}}] \tag{27}$$

Using these definitions, the LS estimate can now be expressed as:

$$\boldsymbol{\theta}_{\mathrm{LS}} = \mathbf{R}_n^{-1}\mathbf{p}_n \tag{28}$$

If the regressors ($\mathbf{x}_k = [x_k^1 \ldots x_k^p]^{\mathrm{T}}$) and output ($y_k$) measurements are subject to zero mean random noise, ie,

$$\tilde{x}^i = x_k^i + n_k^i \quad \text{and} \quad \tilde{y}_k = y_k + w_k, \quad \mathbf{n}_k = [n_k^1 \ldots n_k^p]^{\mathrm{T}} \tag{29}$$

where

$$E[n_k^i] = 0, \quad E[(n_k^i)^2] = v_i, \quad E[w_k] = 0, \quad E[w_k^2] = v_y \tag{30}$$

then analysis of the LS estimate shows that its expected value is given by

$$E[\tilde{\boldsymbol{\theta}}_{\mathrm{LS}}] = [\mathbf{R} + \mathbf{C}_{\mathbf{XX}}]^{-1}(\mathbf{p} + \mathbf{c}_{\mathbf{Xy}}) \tag{31}$$

where

$$E[\mathbf{n}_k \mathbf{n}_k^{\mathrm{T}}] = \mathbf{C}_{\mathbf{XX}}, \quad E[w_k \mathbf{n}_k] = \mathbf{c}_{\mathbf{Xy}} \quad \text{and} \quad E[w_k^2] = c_{\mathbf{yy}} = v_y \tag{32}$$

Thus in this general case, the LS estimate is strongly biased. If consistent estimates of $\mathbf{C}_{\mathbf{XX}}$ and $\mathbf{c}_{\mathbf{Xy}}$ can be obtained *a priori*, unbiased and consistent parameter estimates can be computed as

$$\tilde{\boldsymbol{\theta}}_{\mathrm{CLS}}{}^{*} = [\tilde{\mathbf{R}}_n - \tilde{\mathbf{C}}_{\mathbf{XX}}]^{-1}(\tilde{\mathbf{p}}_n - \tilde{\mathbf{c}}_{\mathbf{Xy}}) \tag{33}$$

This is referred to as Compensated Least Squares (CLS; Stoica and Soderstrom, 1982). Note '$\sim$' is used to indicate that the values in question are computed from the noisy data $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{y}}$.

If the noise sequences on each regressor and output are independent zero mean sequences with common variance v then

$$\mathbf{C}_{\mathbf{XX}} \to v\mathbf{I}_p, \quad \mathbf{c}_{\mathbf{Xy}} \to 0 \tag{34}$$

and

$$E[\tilde{\boldsymbol{\theta}}_{\mathrm{LS}}] = [\mathbf{R} + v\mathbf{I}_p]^{-1}\mathbf{p} \tag{35}$$

Even in this situation the LS estimate is asymptotically biased:

$$\boldsymbol{\theta}_{\mathrm{bias}} = E[\tilde{\boldsymbol{\theta}}_{\mathrm{LS}}] - \boldsymbol{\theta}_{\mathrm{LS}} = [\mathbf{R} + v\mathbf{I}_p]^{-1}\mathbf{p} - \mathbf{R}^{-1}\mathbf{p} = -v[\mathbf{R} + v\mathbf{I}_p]^{-1}\boldsymbol{\theta}_{\mathrm{LS}} \tag{36}$$

In fact the only situation where LS produces unbiased estimates is when zero mean noise is present on the output only and the regressors are noise free. However, an alternative approach known as TLS can be used to obtain unbiased estimate of $\boldsymbol{\theta}$ when noise is present on the regressors.

2.4 Total Least Squares (TLS)

TLS seeks to achieve the best fit to the linear model (18) when noise is present on both $\mathbf{X}$ and $\mathbf{y}$. Thus if (18) represents the true model, then the matrix equation formed by the noisy data ($\tilde{\mathbf{X}}$ and $\tilde{\mathbf{y}}$):

$$\tilde{\mathbf{X}}\boldsymbol{\theta} \approx \tilde{\mathbf{y}} \tag{37}$$

will be incompatible. The LS solution can be interpreted as finding the smallest perturbation of $\tilde{\mathbf{y}}$, $\Delta\mathbf{y}$, that makes the equation compatible. Thus the LS solution corresponds to

$$\Delta\mathbf{y}^{\mathrm{LS}} = arg\{\min_{\Delta\mathbf{y}} ||\Delta\mathbf{y}||_2\} \text{ subject to the constraint } \tilde{\mathbf{X}}\boldsymbol{\theta} = \tilde{\mathbf{y}} - \Delta\mathbf{y} \tag{38}$$

Any solution of $\tilde{\mathbf{X}}\tilde{\boldsymbol{\theta}}_{\mathrm{LS}} = \tilde{\mathbf{y}} - \Delta\mathbf{y}^{\mathrm{LS}}$ is then an LS estimate of the parameters.

The objective of TLS on the other hand is to find the minimal perturbation of both $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{y}}$ to make the equation compatible. Thus the TLS problem can be formulated as follows: find $\Delta\mathbf{X}$ and $\Delta\mathbf{y}$ such that

$$\Delta\mathbf{X}^{\mathrm{TLS}}, \quad \Delta\mathbf{y}^{\mathrm{TLS}} = arg\left\{\min_{\Delta\mathbf{X},\,\Delta\mathbf{y}} ||[\Delta\mathbf{X}\ \Delta\mathbf{y}]||_F\right\} \tag{39}$$

subject to the constraint

$$[\tilde{\mathbf{X}} - \Delta\mathbf{X}]\boldsymbol{\theta} = \tilde{\mathbf{y}} - \Delta\mathbf{y} \tag{40}$$

Any $\boldsymbol{\theta}$ satisfying

$$[\tilde{\mathbf{X}} - \Delta\mathbf{X}^{\mathrm{TLS}}]\boldsymbol{\theta} = \tilde{\mathbf{y}} - \Delta\mathbf{y}^{\mathrm{TLS}} \tag{41}$$

is a TLS parameter estimate.

The TLS estimate can be computed using SVD as follows. Rewriting (37) as

$$[\tilde{\mathbf{X}}\,\tilde{\mathbf{y}}]\begin{bmatrix}\boldsymbol{\theta}\\ -1\end{bmatrix} \approx \mathbf{0} \tag{42}$$

and assuming $rank\,([\tilde{\mathbf{X}}\,\tilde{\mathbf{y}}]) = p+1,$ then finding the TLS solution becomes a problem of determining the minimum change to augmented matrix $[\tilde{\mathbf{X}}\,\tilde{\mathbf{y}}]$ so that its rank is reduced to p making (42) compatible. The required perturbation can be determined by performing a SVD on $[\tilde{\mathbf{X}}\,\tilde{\mathbf{y}}]$. Thus

$$[\tilde{\mathbf{X}}\,\tilde{\mathbf{y}}] = \mathbf{U}_{p+1}\boldsymbol{\Sigma}_{p+1}\mathbf{V}_{p+1}^{\mathrm{T}} = \sum_{i=1}^{p+1}\sigma_i\mathbf{u}_i\mathbf{v}_i^{\mathrm{T}} \tag{43}$$

The minimal TLS correction is

$$[\Delta\mathbf{X}^{\mathrm{TLS}}\ \Delta\mathbf{y}^{\mathrm{TLS}}] = \sigma_{p+1}\mathbf{u}_{p+1}\mathbf{v}_{p+1}^{\mathrm{T}} \tag{44}$$

and the corresponding rank p perturbed matrix is

$$[\mathbf{X}^{\mathrm{TLS}}\ \mathbf{y}^{\mathrm{TLS}}] = \sum_{i=1}^{p}\sigma_i\mathbf{u}_i\mathbf{v}_i^{\mathrm{T}} \tag{45}$$

Now the matrix equation

$$[\mathbf{X}^{\mathrm{TLS}}\ \mathbf{y}^{\mathrm{TLS}}]\begin{bmatrix}\boldsymbol{\theta}\\ -1\end{bmatrix} = \mathbf{0} \tag{46}$$

is compatible and any $\boldsymbol{\theta}$ satisfying it is a TLS solution. However, since $\mathbf{v}_{p+1}$ is the only vector in the null space of $[\mathbf{X}^{\mathrm{TLS}}\ \mathbf{y}^{\mathrm{TLS}}]$, the TLS solution is unique and is given by $\mathbf{v}_{p+1}$ scaled so that its last component is -1. Thus

$$\begin{bmatrix}\boldsymbol{\theta}_{\mathrm{TLS}}\\ -1\end{bmatrix} = -\left[\frac{1}{v_{p+1,\,p+1}}\right]\mathbf{v}_{p+1} \tag{47}$$

By exploiting the fact that only the minor singular vector is needed for determining the TLS solution, an efficient partial SVD algorithm can be used instead of full SVD (Van Huffel and Vandewalle, 1991).

The significance of the TLS solution is that when the perturbations $[\Delta\mathbf{X}\ \Delta\mathbf{y}]$ from the true data $[\mathbf{X}\ \mathbf{y}]$, ie,

$$[\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}] = [\mathbf{X}\ \mathbf{y}] + [\Delta\mathbf{X}\ \Delta\mathbf{y}] \tag{48}$$

are due to independent zero mean noise sequences with common variance, then TLS produces unbiased and consistent estimates of $\boldsymbol{\theta}_{\mathrm{LS}}$ (Van Huffel and Vandewalle, 1991).

It should be noted that the covariance matrix of the TLS estimate is always larger that the covariance matrix of the LS estimate and therefore for low noise levels LS may give better overall performance (ie, a lower mean squared estimation). The covariance of TLS estimates is approximately given by

$$E[(\boldsymbol{\theta}_{\mathrm{TLS}} - \boldsymbol{\theta})^2] \approx (1 + \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta})\frac{v\mathbf{R}^{-1}}{n} \tag{49}$$

where as the covariance for the LS estimate is given by

$$E[(\boldsymbol{\theta}_{\mathrm{LS}} - \boldsymbol{\theta})^2] = \frac{v\mathbf{R}^{-1}}{n} \tag{50}$$

2.5 Generalized total least squares (GTLS)

Unfortunately the TLS solution only gives unbiased estimates when the noise on $[\mathbf{X}\ \mathbf{y}]$ has equal variance and is mutually independent, ie,

$$E\left[\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]^{\mathrm{T}}\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]\right] = \mathbf{C} = \begin{bmatrix} \mathbf{C}_{\mathbf{XX}} & \mathbf{c}_{\mathbf{Xy}} \\ \mathbf{c}_{\mathbf{Xy}}^{\mathrm{T}} & c_{\mathbf{yy}} \end{bmatrix} = v\mathbf{I}_{p+1} \tag{51}$$

This is quite a restrictive requirement. CLS can be used if $\mathbf{C}$ is known or can be consistently estimated (Stoica and Soderstrom, 1982). Alternatively, if the structure of $\mathbf{C}$ is known up to a factor of proportionality, ie, $\mathbf{C} = v\mathbf{C}_0$, then the augmented matrix $[\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}]$ can be transformed by a weighting matrix $\mathbf{W}$ so that the resulting noise covariance matrix meets the unbiasedness requirement. This modified TLS algorithm is known as GTLS and is computed as follows.

Given

$$\left[\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}\right]\begin{bmatrix} \boldsymbol{\theta} \\ -1 \end{bmatrix} \approx 0 \tag{52}$$

with noise covariance matrix $\mathbf{C} = v\mathbf{C}_0$ where v is unknown, then an unbiased estimate of $\boldsymbol{\theta}$ is obtained by applying TLS to the weighted matrix equation

$$\left[\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}\right]\mathbf{W}^{-1}\begin{bmatrix} \boldsymbol{\theta}_w \\ -1 \end{bmatrix} \approx \mathbf{0} \tag{53}$$

and then computing

$$\mathbf{v}^w = \mathbf{W}^{-1}\begin{bmatrix} \boldsymbol{\theta}_w \\ -1 \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{\theta}_{\mathrm{GTLS}} \\ -1 \end{bmatrix} = -\frac{\mathbf{v}^w}{v^w_{p+1,\,p+1}} \tag{54}$$

Since

$$E\left[\left[\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]\mathbf{W}^{-1}\right]^{\mathrm{T}}\left[\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]\mathbf{W}^{-1}\right]\right] = v[\mathbf{W}^{-1}]^{\mathrm{T}}\mathbf{C}_0\mathbf{W}^{-1} \tag{55}$$

and we require $[\mathbf{W}^{-1}]^{\mathrm{T}}\mathbf{C}_0\mathbf{W}^{-1} = \mathbf{I}_{p+1}$ then $\mathbf{C}_0 = \mathbf{W}^{\mathrm{T}}\mathbf{W}$. Therefore the required weighting matrix $\mathbf{W}$ is the Cholesky decomposition of $\mathbf{C}_0$ (or the square root of $\mathbf{C}_0$).

In practice, this direct approach to GTLS is not used because of potential ill-conditioning arising from $\mathbf{W}^{-1}$. Instead the solution can be determined in a robust fashion by taking the generalized SVD (*gsvd*; Van Huffel and Vandewalle, 1989) of the matrix pair $[\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}]$ and $\mathbf{W}$, and determining $\boldsymbol{\theta}_{\mathrm{GTLS}}$ from the generalized singular vector associated with the smallest generalized singular value, ie,

$$gsvd([\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}], \mathbf{W}) \rightarrow \begin{cases} [\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}] = \mathbf{U}\boldsymbol{\Sigma}_{\mathbf{Xy}}\mathbf{G}^{-1} \\ \mathbf{W} = \mathbf{V}\boldsymbol{\Sigma}_{\mathbf{W}}\mathbf{G}^{-1} \\ \boldsymbol{\Sigma}^2 = \boldsymbol{\Sigma}_{\mathbf{Xy}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{Xy}}[\boldsymbol{\Sigma}_{\mathbf{W}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{W}}]^{-1} \end{cases} \tag{56}$$

where $\boldsymbol{\Sigma}_{\mathbf{W}} \in \Re^{(p+1)\times(p+1)}$ and $\boldsymbol{\Sigma}_{\mathbf{Xy}} \in \Re^{(p+1)\times(p+1)}$ are diagonal matrices with singular values on their diagonals and $\boldsymbol{\Sigma}$ contains the generalized singular values. The corresponding generalized singular vectors are contained in the matrix $\mathbf{G} = [\mathbf{g}_1 \ldots \mathbf{g}_{p+1}]$. Therefore

$$\boldsymbol{\theta}_{\mathrm{GTLS}} = -\frac{\mathbf{g}_{p+1}}{g_{p+1,\,p+1}} \tag{57}$$

where $\mathbf{g}_{p+1}$ is the vector associated with the smallest generalized singular value. The corresponding variance on the GTLS parameter estimates is approximately given by

$$E[(\boldsymbol{\theta}_{\mathrm{GTLS}} - \boldsymbol{\theta})^2] \approx \left[\boldsymbol{\theta}^T - 1\right]\mathbf{C}\begin{bmatrix} \boldsymbol{\theta} \\ -1 \end{bmatrix}\frac{v\mathbf{R}^{-1}}{n} \tag{58}$$

3. The difference equation TTP technique

3.1 Basic principles

For a given sample rate, T_s, the thermocouples constituting a TTP can be modelled as first-order difference equations of the form

$$T_1^k = a_1 T_1^{k-1} + b_1 T_g^{k-1} \tag{59}$$

$$T_2^k = a_2 T_2^{k-1} + b_2 T_g^{k-1} \tag{60}$$

These can be related to the continuous time equations describing the TTP (4 and 5 above) under the assumption of ZOH on the input signal (T_g), ie,

$$a_1 = \exp\left(-\frac{T_s}{\tau_i}\right) \text{ and } b_i = 1 - a_i, \quad i = 1, 2 \tag{61}$$

While ZOH is clearly not true for a continuously changing gas temperature, it becomes a valid approximation provided the system is sufficiently over-sampled, ie, $f_s \gg 2f_{g(\max)}$ where $f_{g(\max)}$ is the bandwidth of T_g and f_s is the sampling frequency. Thus, if the parameters of the discrete model equations can be determined, the thermocouple time constants can then be estimated as

$$\tau_i = -\frac{T_s}{\ln(a_i)} = -\frac{T_s}{\ln(1-b_i)}, \quad i = 1, 2 \tag{62}$$

Unknown signal T_g^{k-1} can be eliminated from the simultaneous equations formed by (59) and (60) to give the difference equation TTP model as

$$\frac{1}{b_1} T_1^k - \frac{a_1}{b_1} T_1^{k-1} = \frac{1}{b_2} T_2^k - \frac{a_2}{b_2} T_2^{k-1} \tag{63}$$

This can be rewritten with T_2 considered as the response variable and T_1 as the input variable, ie,

$$T_2^k = a_2 T_2^{k-1} + \frac{b_2}{b_1} T_1^k - \frac{b_2}{b_1} a_1 T_1^{k-1} \tag{64}$$

or

$$T_2^k = a_2 T_2^{k-1} + \beta T_1^k - \beta a_l T_1^{k-1} \tag{65}$$

where $\beta = b_2/b_1$ is introduced here to simplify the expression.

3.2 Estimating the difference equation TTP model parameters

Expressing the TTP model (Equation 65) in vector form

$$\left[T_2^k\right] = \left[\begin{array}{ccc} T_2^{k-1} & T_1^k & -T_1^{k-1} \end{array}\right] \left[\begin{array}{c} a_2 \\ \beta \\ \beta a_1 \end{array}\right] \tag{66}$$

it can be seen that given $n \geq 4$ samples of T_1 and T_2 unknowns a_2, β and $\boldsymbol{\theta}_3 (= \beta a_1)$ can be uniquely determined using an appropriate LS estimation algorithm. Here,

$$\mathbf{x}_k^{\mathrm{T}} = \left[\begin{array}{ccc} T_2^{k-1} & T_1^k & -T_1^{k-1} \end{array}\right], \quad y_k = T_2^k \text{ and } \boldsymbol{\theta} = \left[\begin{array}{c} a_2 \\ \beta \\ \theta_3 \end{array}\right] \tag{67}$$

Parameter a_1 can then be estimated as $a_1 = \theta_3/\beta$.

3.2.1 Bias on estimates: If the measured temperatures T_1 and T_2 are subject to zero-mean identically distributed white noise with variances v_1 and v_2, respectively,

then the expected correlation matrix and cross-correlation vector are

$$E[\tilde{\mathbf{R}}_n] = \mathbf{R} + \begin{bmatrix} v_2 & 0 & 0 \\ 0 & v_1 & 0 \\ 0 & 0 & v_1 \end{bmatrix}, \quad E[\tilde{\mathbf{p}}_n] = \mathbf{p} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{68}$$

Thus the LS estimate will be biased, but unbiased estimates can be expected with TLS if $v_1{=}v_2$ and with GTLS if the ratio of the noise variances, $\phi{=}v_1/v_2$, is known. Note that for this formulation the noise covariance matrix as defined in (51) is given by

$$\mathbf{C} = diag(v_2, v_1, v_1, v_2) = v_2 diag(1, \phi, \phi, 1) \tag{69}$$

It is not possible uniquely to identify b_1 and b_2. Only their ratio β can be determined unless additional information is available. In the case of unity gain thermocouples, we have the additional constraint that in steady-state

$$T_1 = T_2 = T_g \tag{70}$$

hence from Equations (59) and (60)

$$b_1 = 1 - a_1, \; b_2 = 1 - a_2 \tag{71}$$

Thus, we do not have to compute b_1 and b_2; rather we have to solve Equation (66) subject to the constraint that

$$\beta = \frac{1 - a_2}{1 - a_1} \tag{72}$$

When fitting data from a truly linear system to this model (ie, by estimating a_2, β and $\boldsymbol{\theta}_3$ and computing $a_1 = \boldsymbol{\theta}_3/\beta$), the constraint will implicitly be satisfied. However, in the presence of noise or system non-linearity, the extra degree of freedom provided by the third parameter in the effectively two-parameter model can lead to poor estimates (over-fitting). This can be assessed by monitoring the degree of violation of Equation (72) by the estimated parameters.

The identification task can be reduced to a two-parameter problem by substituting Equation (72) into (66) giving

$$T_2^k = a_2 T_2^{k-1} + \left(\frac{1 - a_2}{1 - a_1}\right) T_1^k - a_1 \left(\frac{1 - a_2}{1 - a_1}\right) T_1^{k-1} \tag{73}$$

This is no longer a linear in the parameter model and must be solved by non-linear methods. This typically involves minimizing a mean squared estimation cost function of the form

$$J(a_1, a_2) = E_n \left[\left(T_2^k - a_2 T_2^{k-1} - \left(\frac{1 - a_2}{1 - a_1}\right) T_1^k + a_1 \left(\frac{1 - a_2}{1 - a_1}\right) T_1^{k-1} \right)^2 \right] \tag{74}$$

where

$$E_n[x_k] = \frac{1}{n}\sum_{i=1}^{n} x_k \tag{75}$$

This cost function is highly non-linear with a singularity at $a_1 = 1$. While there is only a single global minimum, minimization by iterative gradient-based methods is poorly conditioned leading to slow convergence and numerical issues. In addition, minimization in the presence of noise is biased, but there is no systematic approach for dealing with this in a non-linear setting. Consequently, alternative linear formulations are needed.

4. λ-formulations

4.1 Three-parameter λ-formulation

In Hung *et al*. (2003), the authors propose a three-parameter λ-formulation where Equation (73) is written as

$$T_2^k = \lambda_1 T_2^{k-1} + \lambda_2 T_1^k + \lambda_3 T_1^{k-1} \tag{76}$$

which corresponds to identifying the discrete-time model given in Figure 4. Following identification of the linear model parameters, λ_1, λ_2, λ_3 the desired coefficients a_1 and a_2 can be determined according to

$$a_1 = -\frac{\lambda_3}{\lambda_2}, \quad a_2 = \lambda_1 \tag{77}$$

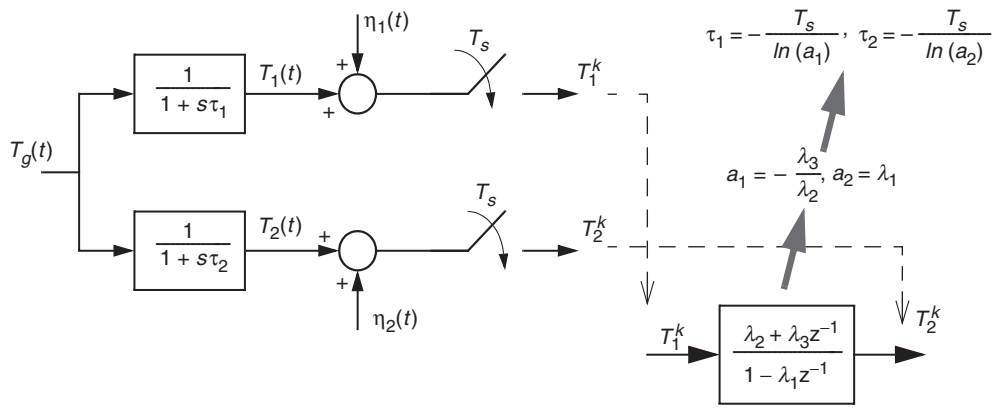

Figure 4 Discrete-time TTP formulation

This formulation is essentially equivalent to that given in Equation (66) and therefore has similar issues. The noise covariance matrix, $\mathbf{C}_{\lambda_3}$ is as defined in Equation (69). However, an advantage of this parameterization is that the constraint on the extra degree of freedom can be concisely expressed as

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{78}$$

4.2 Two-parameter λ-formulations

Significantly, substitution of the constraint as defined by (78) into the identification model in (76) reduces parameter estimation to a two-dimensional linear optimization problem, the exact form of which depends on which of the three unknown parameters is eliminated using Equation (78). For example, if we eliminate λ_3 by substituting

$$\lambda_3 = 1 - \lambda_1 - \lambda_2 \tag{79}$$

then (76) becomes

$$T_2^k = \lambda_1 T_2^{k-1} + \lambda_2 T_1^k + (1 - \lambda_1 - \lambda_2) T_1^{k-1} \tag{80}$$

Collecting terms yields

$$T_2^k - T_1^{k-1} = \lambda_1 (T_2^{k-1} - T_1^{k-1}) + \lambda_2 (T_1^k - T_1^{k-1}) \tag{81}$$

$$\left[T_2^k - T_1^{k-1} \right] = \begin{bmatrix} T_2^{k-1} - T_1^{k-1} & T_1^k - T_1^{k-1} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \tag{82}$$

4.2.1 Bias on estimates: The form of the bias in the LS solution is more complex in this formulation because of the multiple occurrences of T_1 and T_2 in the regressor and output, violating the assumptions on noise required by LS and TLS. If we assume zero mean noise with variances v_1 and v_2 as before, then the resulting estimates of the correlation matrix and cross-correlation vector are

$$E[\tilde{\mathbf{R}}_n] = \mathbf{R} + \begin{bmatrix} v_1 + v_2 & v_1 \\ v_1 & 2v_1 \end{bmatrix}, \quad E[\tilde{\mathbf{p}}_n] = \mathbf{p} + \begin{bmatrix} v_1 \\ v_1 \end{bmatrix} \tag{83}$$

and the overall noise covariance matrix as defined in (51) is given by

$$\mathbf{C}_{\lambda_{2a}} = v_2 \begin{bmatrix} \phi + 1 & \phi & \phi \\ \phi & 2\phi & \phi \\ \phi & \phi & \phi + 1 \end{bmatrix} \tag{84}$$

Thus, both LS and TLS will produce biased estimates. Unbiased estimates can only be achieved if estimates of the variances v_1 and v_2 are available (CLS) or their ratio ϕ is known (GTLS).

The other possibilities are

$$\left[T_2^k - T_1^k\right] = \left[\begin{array}{cc} T_2^{k-1} - T_1^k & T_1^{k-1} - T_1^k \end{array}\right]\left[\begin{array}{c} \lambda_1 \\ \lambda_3 \end{array}\right] \tag{85}$$

and

$$\left[T_2^k - T_2^{k-1}\right] = \left[\begin{array}{cc} T_1^k - T_2^{k-1} & T_1^{k-1} - T_2^{k-1} \end{array}\right]\left[\begin{array}{c} \lambda_2 \\ \lambda_3 \end{array}\right] \tag{86}$$

The corresponding noise covariance matrix in each case is

$$\mathbf{C}_{\lambda_{2b}} = v_2\left[\begin{array}{ccc} \phi+1 & \phi & \phi \\ \phi & 2\phi & \phi \\ \phi & \phi & \phi+1 \end{array}\right] \tag{87}$$

$$\mathbf{C}_{\lambda_{2c}} = v_2\left[\begin{array}{ccc} \phi+1 & 1 & 1 \\ 1 & \phi+1 & 1 \\ 1 & 1 & 2 \end{array}\right] \tag{88}$$

Since the different formulations are related algebraically in a manner that is independent of the measurements, eg,

$$\left[\begin{array}{c} \lambda_2 \\ \lambda_3 \end{array}\right] = \left[\begin{array}{c} 0 \\ 1 \end{array}\right] + \left[\begin{array}{cc} 0 & 1 \\ -1 & -1 \end{array}\right]\left[\begin{array}{c} \lambda_1 \\ \lambda_2 \end{array}\right] \tag{89}$$

it follows that the statistical properties are not affected by these transformations. Hence the bias and covariance of the time constant estimates obtained is the same for all the two-parameter formulations. Indeed, when conventional LS is used, the resulting bias and covariance will be identical to that obtained from minimization of the non-linear two-parameter cost function defined by (74). Note, also that in general the bias and variance obtained with the two-parameter formulations will be less than those obtained with the three-parameter formulation defined in (76), since the extra degree of freedom in this formulation will result in increased variance. Clearly, the linear formulations are preferred to the non-linear formulation as unbiased estimates can be obtained using GTLS. The three-parameter formulation is of value when we are restricted to using TLS.

4.3 Sensitivity to noise variance ratio estimate

Which of the two-parameter formulations is of most value is an open question. The appropriate choice may depend on factors such as problem conditioning, and sensitivity to the estimate of noise variance ratio, ϕ. Interestingly, $\mathbf{C}_{\lambda 2a} = \mathbf{C}_{\lambda 2b}$ while $\mathbf{C}_{\lambda 2c}$ is substantially different and contains fewer elements that are functions of ϕ. Consequently, it might be expected that the latter will be less sensitive to errors in ϕ. This can be quantified in terms of the Frobenius norm as follows:

$$||\mathbf{C}_{\lambda 2a}||_F = \sqrt{12\phi^2 + 4\phi + 2} \tag{90}$$

$$||\mathbf{C}_{\lambda 2c}||_F = \sqrt{2\phi^2 + 4\phi + 12} \tag{91}$$

hence $\|\mathbf{C}_{\lambda 2a}\|_F > \|\mathbf{C}_{\lambda 2c}\|_F$ when $\phi > 1$. Measuring sensitivity as the change in Frobenius norm due to a change in ϕ gives

$$\Delta||\mathbf{C}_{\lambda 2a}||_F = \frac{12\phi + 2}{||\mathbf{C}_{\lambda 2a}||_F}\Delta\phi \tag{92}$$

and

$$\Delta||\mathbf{C}_{\lambda 2c}||_F = \frac{2\phi + 2}{||\mathbf{C}_{\lambda 2c}||_F}\Delta\phi \tag{93}$$

Defining the sensitivity in terms of percentages, we have

$$S^{\phi}_{\mathbf{C}_{\lambda 2a}} = \frac{\%\Delta||\mathbf{C}_{\lambda 2a}||_F}{\%\Delta\phi} = \frac{12\phi^2 + 2\phi}{||\mathbf{C}_{\lambda 2a}||_F^2} = \frac{6\phi^2 + \phi}{6\phi^2 + 2\phi + 1} \tag{94}$$

and

$$S^{\phi}_{\mathbf{C}_{\lambda 2c}} = \frac{\%\Delta||\mathbf{C}_{\lambda 2c}||_F}{\%\Delta\phi} = \frac{2\phi^2 + 2\phi}{||\mathbf{C}_{\lambda 2c}||_F^2} = \frac{\phi^2 + \phi}{\phi^2 + 2\phi + 6} \tag{95}$$

Plots of $S^{\phi}_{\mathbf{C}_{\lambda 2a}}$ and $S^{\phi}_{\mathbf{C}_{\lambda 2c}}$ as a function of ϕ are given in Figure 5(a). The ratio of the sensitivity functions is plotted in Figure 5(b). Clearly, $\mathbf{C}_{\lambda 2a}$ is more sensitive than $\mathbf{C}_{\lambda 2c}$ to errors in ϕ for all $\phi > 0$. The maximum difference occurs when $\phi = 1$ at which point $S^{\phi}_{\mathbf{C}_{\lambda 2a}}$ is 3.5 times $S^{\phi}_{C_{\lambda 2c}}$. In relative terms $\mathbf{C}_{\lambda 2a}$ is 6 to 7 times more sensitive than $\mathbf{C}_{\lambda 2a}$ to errors in ϕ for $\phi < 0.5$.

However, this difference in sensitivity to errors in ϕ does not translate to a corresponding difference in sensitivity of the estimated time constants. Figure 6 shows the sensitivity of time constant estimates to errors in ϕ for the multi-tone random signal benchmark. The two-parameter formulations have identical sensitivity to errors in ϕ and are less sensitive that the three-parameter formulation for all values of ϕ.

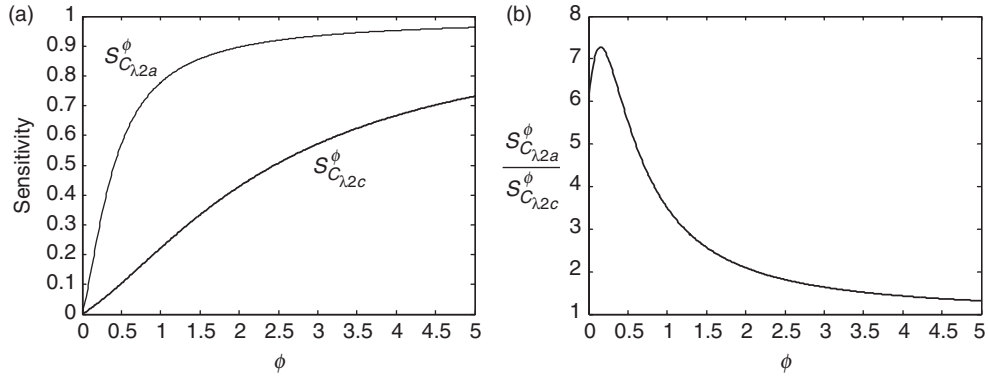

Figure 5 Plots comparing the sensitivity of the two-parameter λ formulation covariance matrices to errors in the noise variance ratio: (a) Sensitivity of $C_{\lambda 2a}$ and $C_{\lambda 2c}$; (b) Relative sensitivity

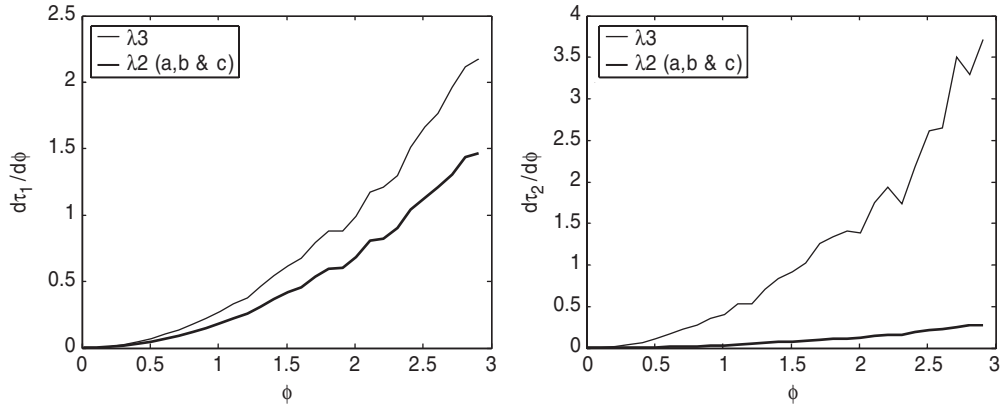

Figure 6 Sensitivity of time constant estimates to errors in ϕ

4.4 Ill-conditioning of the estimation problems

The issue of ill-conditioning also depends on the how it is measured. For conventional LS a good indicator is the condition number of the data covariance matrix (24). When solving the problem using GTLS, the conditioning of the augmented data matrix is more important. Since this matrix should be rank deficient by design, the square of the ratio of the largest to the second smallest singular value is a useful measure of conditioning. Another important consideration for GTLS is the difference between the two smallest singular values, as this determines how effectively the noise can be isolated from the data. Clearly, as the SNR falls this becomes a limiting factor in applying TLS and GTLS methods. It also has implications for algorithm convergence.

Table 1 Singular value ratios for the two-tone benchmark with $\phi = 2$

Singular value ratio/ algorithm formulation	Noise level = 2%, 6%,				Noise level = 10%, 30%			
	λ_{2a}	λ_{2b}	λ_{2c}	λ_3	λ_{2a}	λ_{2b}	λ_{2c}	λ_3
GTLS matrix σ_n/σ_{n+1}	4.8	4.8	4.8	4.8	1.3	1.3	1.3	1.3
GTLS matrix σ_1/σ_n	9.3	9.3	9.3	185.8	6.7	6.7	6.7	132.2
TLS matrix σ_1/σ_n	8.2	8.6	11.6	154.8	5.8	6.1	8.3	109.1
LS matrix σ_1/σ_n	5.9	6.2	12.1	137.1	4.1	4.4	8.4	95.1

Table 2 Singular value ratios for the multi-tone random benchmark with $\phi = 2$

Singular value ratio/ algorithm formulation	Noise level = 2%, 4%				Noise level = 10%, 20%			
	λ_{2a}	λ_{2b}	λ_{2c}	λ_3	λ_{2a}	λ_{2b}	λ_{2c}	λ_3
GTLS matrix σ_n/σ_{n+1}	5.1	5.1	5.1	5.0	1.4	1.4	1.4	1.4
GTLS matrix σ_1/σ_n	8.3	8.3	8.3	15.0	6.0	6.0	6.0	10.8
TLS matrix σ_1/σ_n	7.4	7.6	10.3	14.7	5.4	5.5	7.5	10.5
LS matrix σ_1/σ_n	5.3	5.5	10.7	14.9	3.8	4.0	7.7	10.5

Tables 1 and 2 show typical singular value ratios obtained with the benchmark problems for each of the formulations and optimization methods considered when the noise level on thermocouple 1 is 2% and 10%. The corresponding noise levels on thermocouple 2 for each problem, assuming a noise variance ratio of $\phi = 2$, is as indicated in the tables. The condition number is simply the square of the ratio of the largest and smallest singular values and should be small (ideally 1) for a well conditioned problem. The noise separation ratio (NSR), the ratio of the two smallest singular values (row 1 of the tables), should be as large as possible for good performance of TLS and GTLS methods.

As expected, the two-parameter formulations are equivalent as GTLS problems and are better conditioned than the corresponding three-parameter formulation. As an LS or TLS problem, the λ_{2a} formulation is the best conditioned. Thus, this may be the best choice in situations where the more sophisticated optimization procedures cannot be applied.

Increasing the noise level leads to improved problem conditioning but poorer noise separation. The general trend in these parameters is illustrated in Figure 7 for the multi-tone benchmark. As can be seen, while condition numbers are relatively insensitive to increasing noise level, the NSR decays rapidly and is approximately inversely proportional to noise level. This has major implications for the robustness of TLS and GTLS at high noise levels.

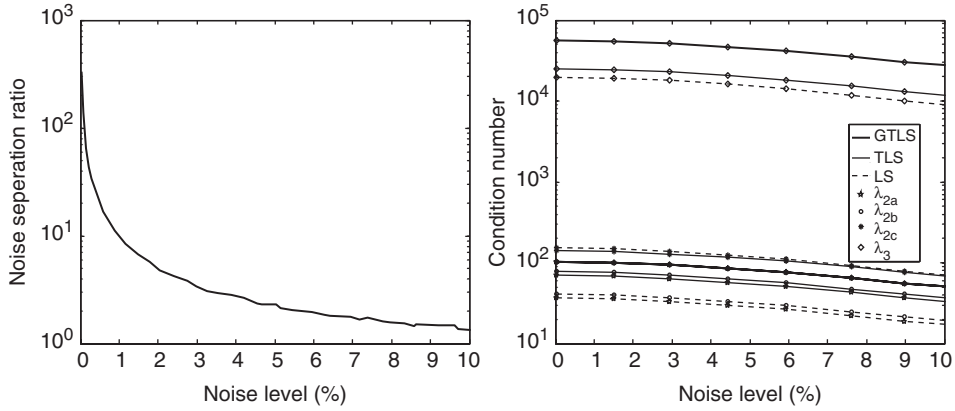

Figure 7 Typical variation in NSR and condition number with noise level

4.5 Determining the noise variance ratio

In the preceding discussions, the only knowledge necessary for consistent unbiased estimates of the time constants (using GTLS) was the ratio of the noise variance, ϕ. This is not a demanding requirement for the TTP application, where experience has shown that probe thermocouples are subject to similar noise environments with the result that $\phi = 1$ can normally be assumed. Furthermore, from the previous discussion, it can be seen that time constant estimation is relatively insensitive to ϕ compared with other factors.

It is possible to estimate the correct noise variance ratio, by parameterizing the GTLS solutions accordingly. The optimum ratio is then the one that produces the minimal perturbation of the data $[\tilde{\mathbf{X}}\ \tilde{\mathbf{y}}]$, ie,

$$\phi = arg\left\{\min_{\phi} \left|\left|\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]_{\phi}^{\mathrm{GTLS}}\right|\right|_F\right\} \tag{96}$$

where

$$\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]_{\phi}^{\mathrm{GTLS}} = arg\left\{\min_{\Delta\mathbf{X},\,\Delta\mathbf{y}} \left|\left|\left[\Delta\mathbf{X}\ \Delta\mathbf{y}\right]\mathbf{W}_{\phi}^{-1}\right|\right|_F\right\} \tag{97}$$

and $\mathbf{W}_{\phi} = \sqrt{\mathbf{C}_{\phi}}$ (ie, the Cholesky factorization of the appropriate noise covariance matrix).

Another possibility is to estimate the noise variance directly by using two pairs of identical thermocouples in each probe (ie, a four-thermocouple probe, two with time

constant τ_1 and two with time constant τ_2). Since the identical thermocouples should be producing identical outputs plus noise, the noise variance can be estimated as

$$v_i = \frac{1}{2n}\sum_{k=1}^{n}[T_{i1}(k) - T_{i2}(k)]^2 \tag{98}$$

Given estimates of both v_1 and v_2, we can choose to use either CLS or GTLS.

5. Results

In this section, results from Monte-Carlo simulations are presented for the TTP measurement of the simulated two-tone and multi-tone benchmark temperature variation signals described earlier. With the TTP time constants, τ_1 and τ_2 selected as 0.02 s and 0.1 s, respectively, and the noise variance ratio, $\phi = 4$, a series of 100-run simulations were performed for noise levels ranging from 0 to 10% (with respect to thermocouple 1) to evaluate the statistical properties of the algorithms. The root-mean squared estimation error (RMSE), bias and variance were computed in each case and are plotted as a function of noise level in Figure 8. In each case, the errors are computed vectorally and expressed as a percentage of the true parameter values. The distribution of parameter estimates and corresponding covariance ellipses (two standard deviations) are displayed in Figure 9 for 1%, 5% and 10% noise levels. These results confirm the theoretical predictions for each algorithm. The bias error dominates in the LS and TLS implementations and the variance error dominates in the GTLS implementations. As the bias is the main contributor to parameter error, it follows that the GTLS algorithms are superior.

As the noise level is increased, GTLS methods deteriorate in performance dramatically because of the rapidly increasing variance. In the multi-tone problem, this deterioration occurs more rapidly with the three-parameter λ-formulation than with the two-parameter formulations, while there is little difference between the methods in the two-tone problem. While not apparent in the two-tone example, formulations with fewer degrees of freedom generally produce parameter estimates with less variance. Consequently the two-parameter λ formulations provide optimum performance.

All the two-parameter λ formulations (λ_{2a}, λ_{2b} and λ_{2c}) yield identical results and are reported as $\lambda 2$ in Figures 8 and 9. It is also worth noting that direct iterative optimization of the non-linear two-parameter cost function $J(a_1, a_2)$ defined by (74) generates identical parameter estimates (and statistical properties) to the two-parameter λ formulations solved using conventional LS ($\lambda 2$–LS).

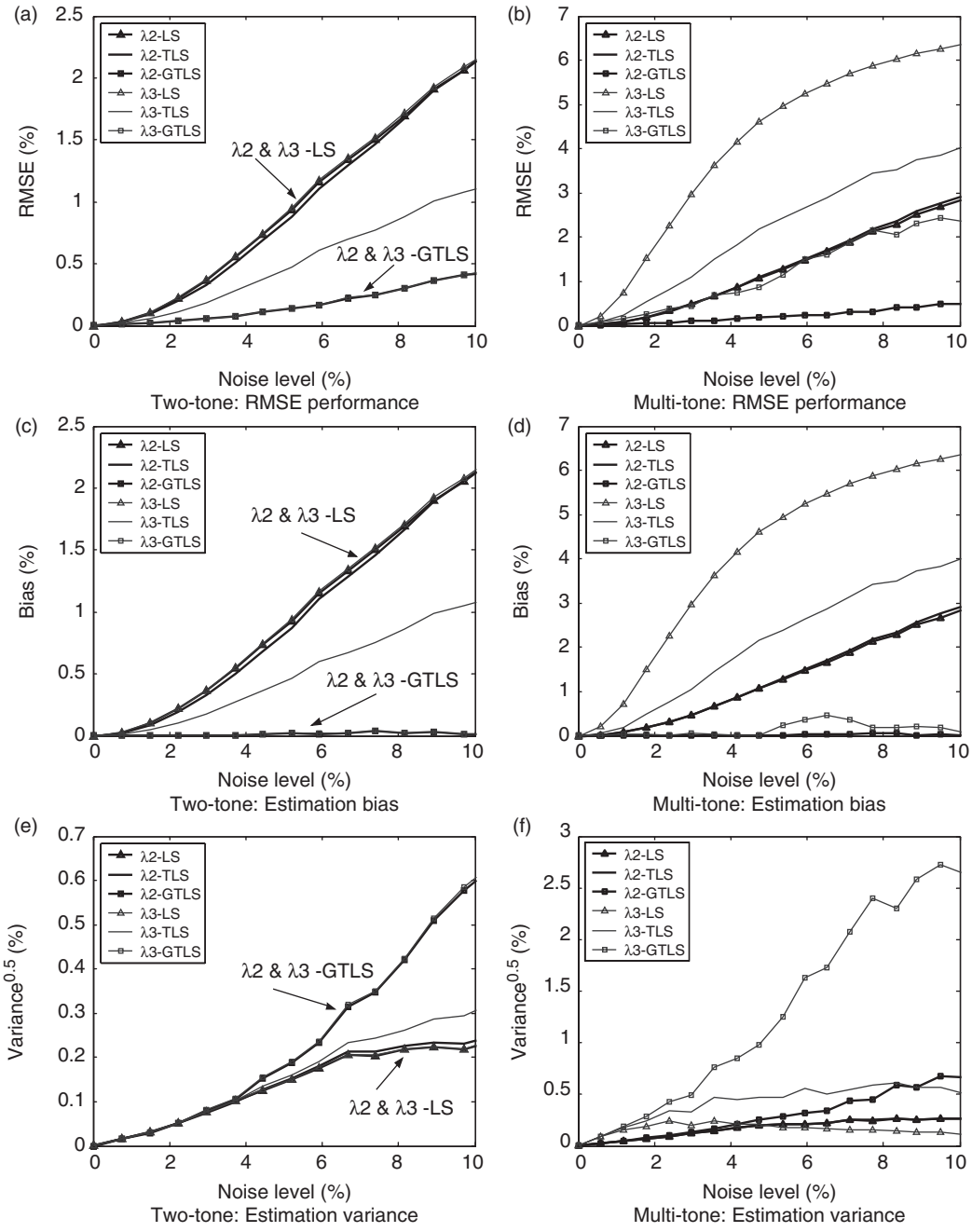

Figure 8 Statistical comparison of algorithms in the presence of measurement noise

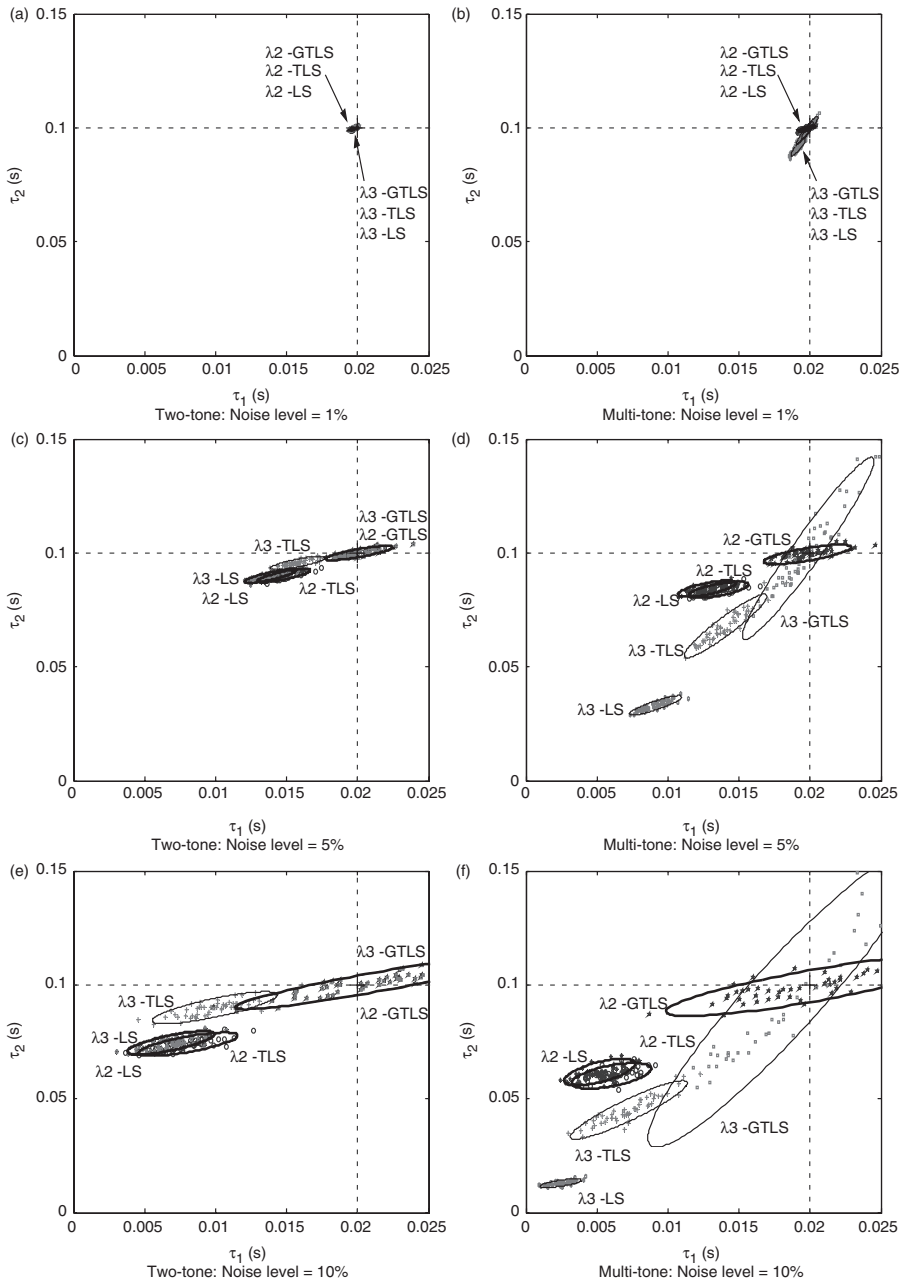

Figure 9 Time constant estimate distributions and covariance ellipses for 1%, 5% and 10% noise levels

6. Conclusions

This paper has explored discrete-time algorithms for TTP characterization. The direct translation from continuous time to discrete time yields a two-parameter non-linear model. However, by appropriate choice of variables, it is possible to reformulate the model as a two-parameter or three-parameter linear optimization problem. This has the advantage of allowing the application of techniques such as TLS and GTLS, which yield unbiased estimates. Several two-parameter formulations have been developed but they all yield equivalent performance, since they are related by invariant transformations of the model. Initial investigation of numerical properties would suggest that they are also equivalent in this respect.

The performance and statistical properties of all the methods presented have been verified through Monte-Carlo tests on two simulated benchmark problems. From these results, it can be deduced that bias is a major source of error in TTP characterization, and consequently methods that can exploit GTLS optimization, or equivalent, to obtain unbiased estimates are of great value. Unfortunately, the variance of estimates obtained with GTLS increases rapidly with noise level limiting the direct application to relative low noise levels. Successful application to higher noise levels is possible in situations where ensemble averaging can be employed.

References

Cambray, P. 1986: Measuring thermocouple time constants: a new method. *Combustion Science and Technology* 45, 221–24.

Forney, L.J. and **Fralick, G.C.** 1994: Two-wire thermocouple: frequency response in constant flow. *Review of Scientific Instruments* 65, 3252–57.

Forney, L.J. and **Fralick, G.C.** 1995a: Three-wire thermocouple: frequency response in constant flow. *Review of Scientific Instruments* 66, 3331–36.

Forney, L.J. and **Fralick, G.C.** 1995b: Multiwire thermocouples in reversing flow. *Review of Scientific Instruments* 66, 5050–54.

Golub, G. and **Van Loan, C.** 1980: An analysis of the total least squares problem. *SIAM Journal of Numerical Analysis* 17, 883–93.

Hung, P.C.F., McLoone, S.F., Irwin, G.W. and **Kee, R.J.** 2002: A novel approach to two wire thermocouple temperature reconstruction. *Proceedings of the Irish Signals and Systems Conference*, Limerick, 193–98.

Hung, P.C.F., McLoone, S.F., Irwin, G.W. and **Kee, R.J.** 2003: A total least squares approach to sensor characterisation. *Proceedings 13th IFAC Symposium on System Identification*, Rotterdam 337–42.

Kee, R.J., O'Reilly, P.G., Fleck, R. and **McEntee, P.T.** 1999: Measurement of exhaust gas temperature in a high performance two-stroke engine. *SAE Transactions Journal of Engines* 107, 983072.

Ljung, L. 1999: *System identification: theory for the user*. Prentice Hall.

O'Reilly, P.G., Kee, R.J., Fleck, R. and **McEntee, P.T.** 2001: Two-wire thermocouples: a nonlinear state estimation approach to temperature reconstruction. *Review of Scientific Instruments* 72, 3449–57.

Pfriem, H. 1936: Zur Messung veränderlicher Temperaturen von Gasen und Flüssigkeiten. *Gen. Ingen* 7, 85–92.

Stoica, P. and **Soderstrom, T.** 1982: Bias correction in least-squares identification. *International Journal of Control* 35, 449–57.

Strahle, W.C. and **Muthukrishman, M.** 1976: Thermocouple time constant measurement by cross power spectrum. *Journal of the American Institute of Aeronautics and Astronautics* 14, 1642–44.

Tagawa, M., **Kato, K.** and **Ohta, Y.** 2003: Response compensation of temperature sensors: frequency-domain estimation of thermal time constants. *Review of Scientific Instruments* 74, 3171–74.

Tagawa, M. and **Ohta, Y.** 1997: Two-thermocouple probe for fluctuating temperature measurement in combustion – rational estimation of mean and fluctuating time constants. *Combustion and Flame* 109, 549–60.

Tagawa, M., **Shimoji, T.** and **Ohta, Y.** 1998: A two-thermocouple probe technique for estimating thermocouple time constants in flows with combustion: *in situ* parameter identification of a first-order lag system. *Review of Scientific Instruments* 69, 3370–78.

Van Huffel, S. and **Vandewalle, J.** 1989: Analysis and properties of the generalised total least squares problem AX=B when some or all columns in A are subject to error. *SIAM Journal of Matrix Analysis and Applications* 10, 294–15.

Van Huffel, **S.** and **Vandewalle**, **J.** 1991: *The total least squares problem: computational aspects and analysis*. SIAM.